FIG. 1

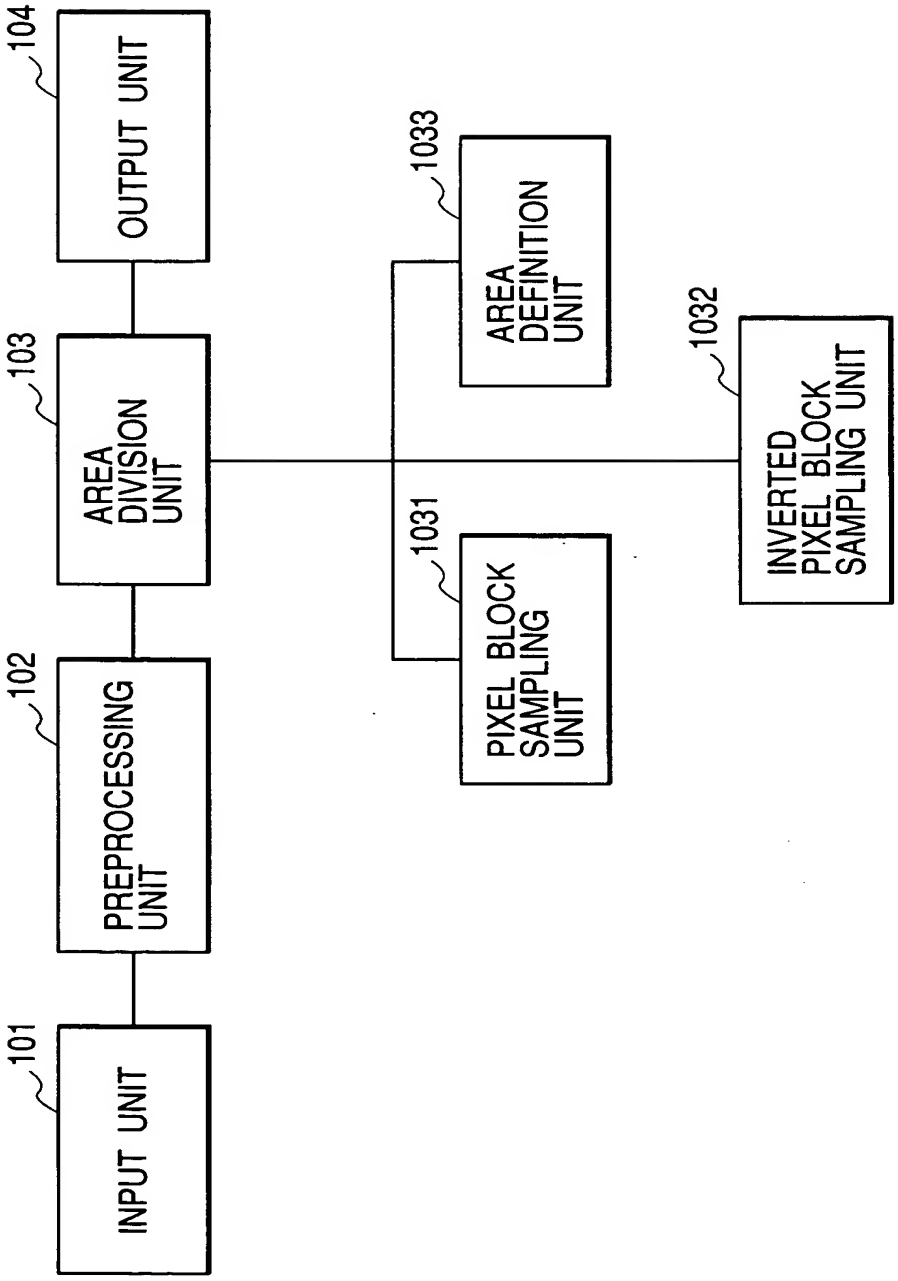

101
INPUT UNIT
102
PREPROCESSING UNIT
103
AREA DIVISION UNIT
104
OUTPUT UNIT
1031
PIXEL BLOCK SAMPLING UNIT
1032
INVERTED PIXEL BLOCK SAMPLING UNIT
1033
AREA DEFINITION UNIT

FIG. 2

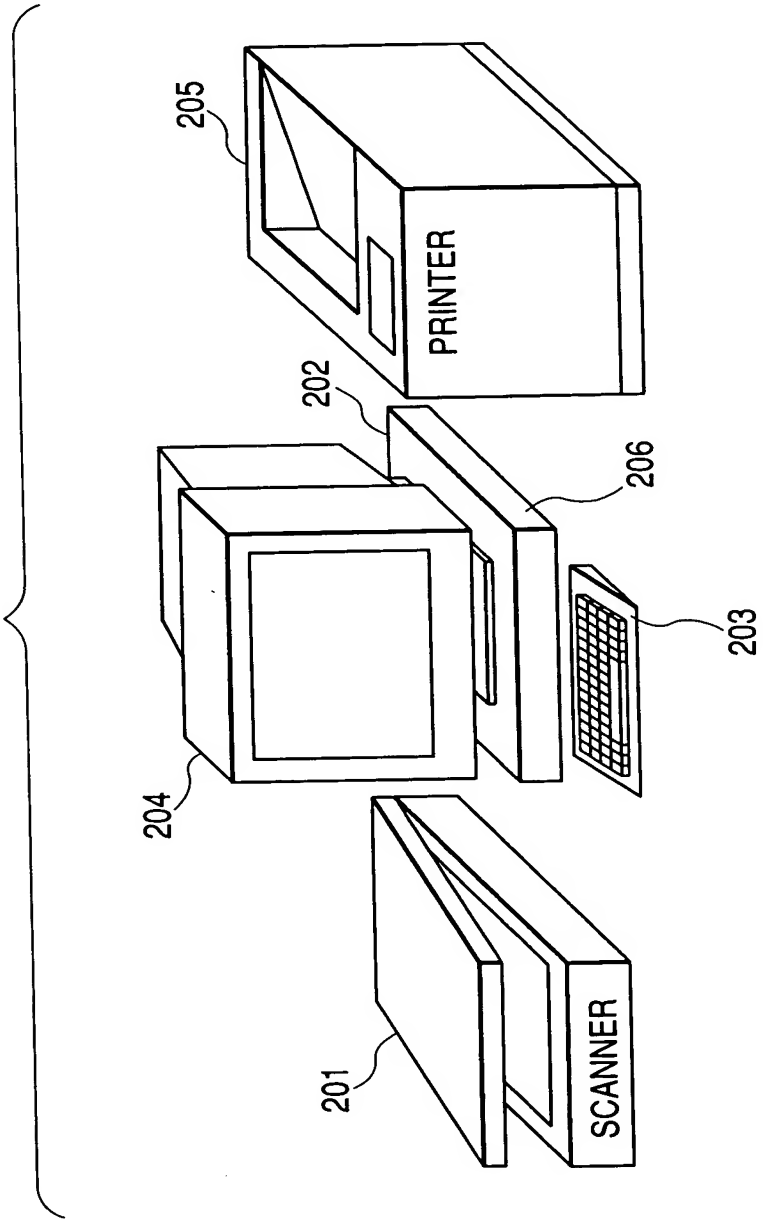


FIG. 3

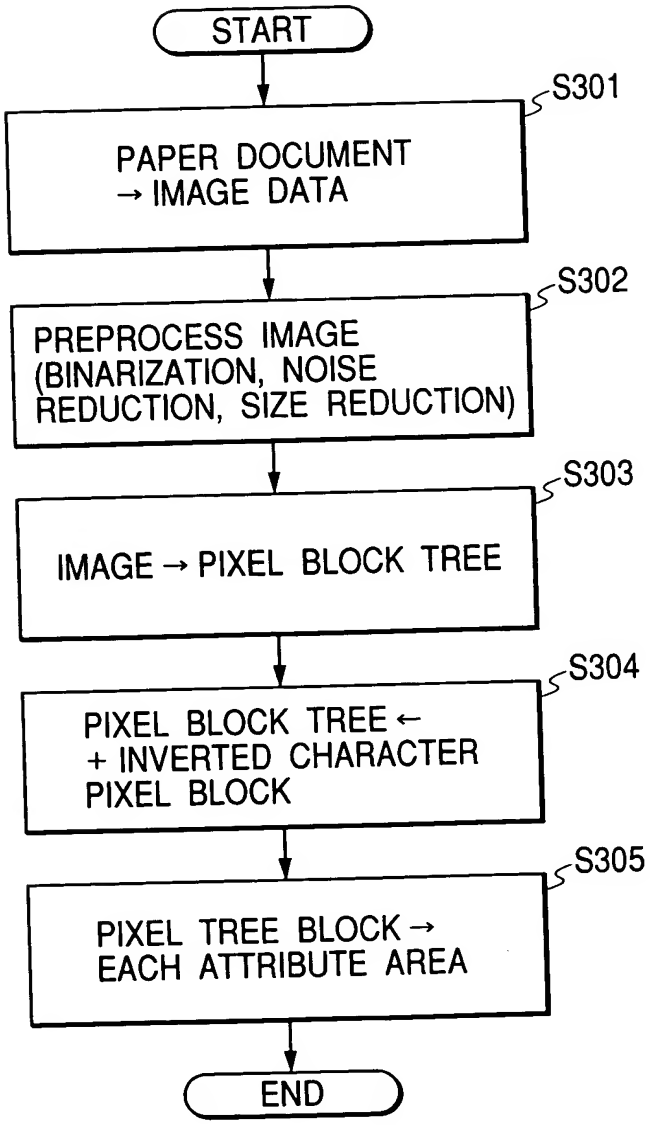

START
S301
PAPER DOCUMENT → IMAGE DATA
S302
PREPROCESS IMAGE (BINARIZATION, NOISE REDUCTION, SIZE REDUCTION)
S303
IMAGE → PIXEL BLOCK TREE
S304
PIXEL BLOCK TREE ← + INVERTED CHARACTER PIXEL BLOCK
S305
PIXEL TREE BLOCK → EACH ATTRIBUTE AREA
END

FIG. 4

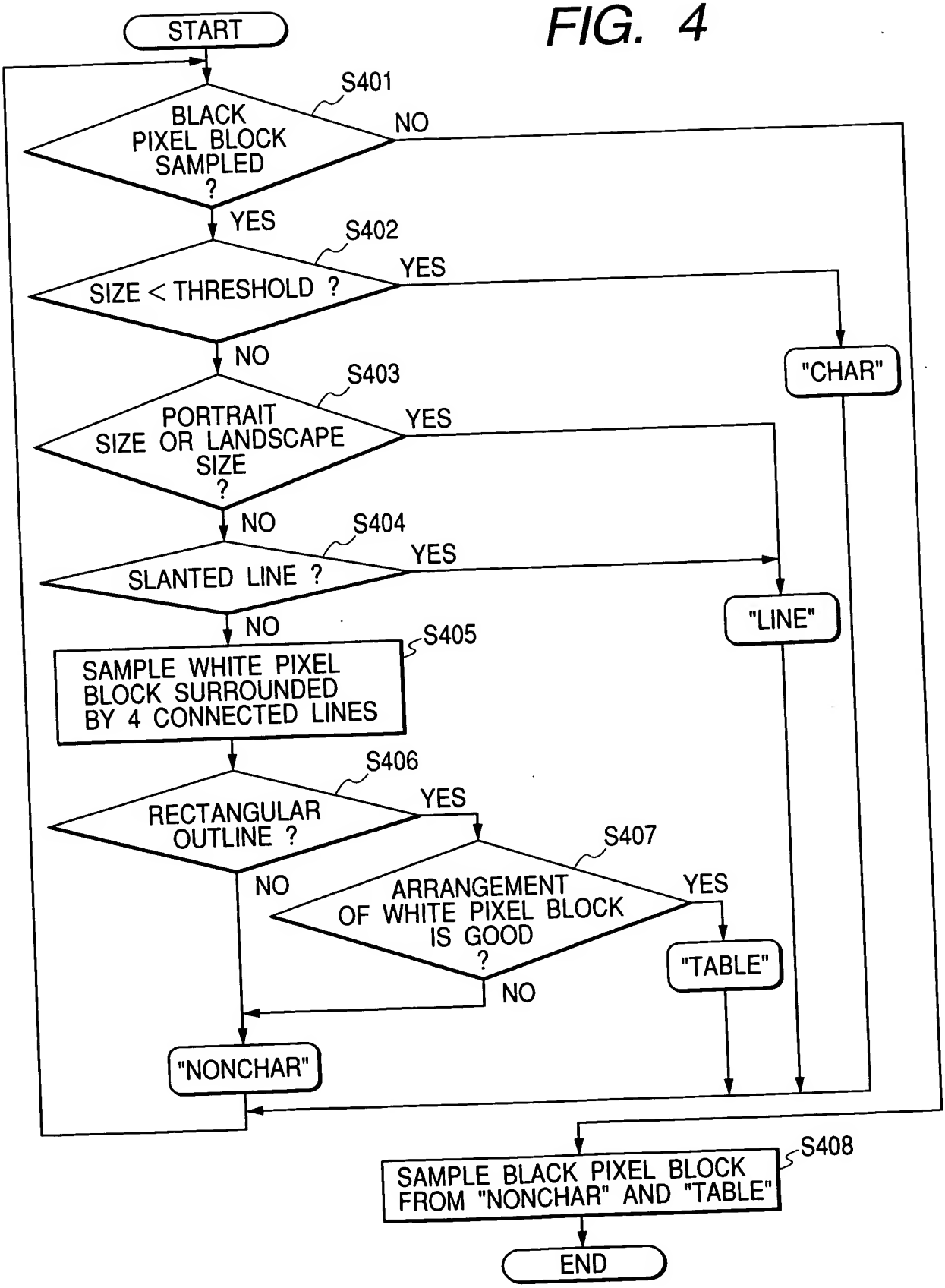

START
S401
BLACK PIXEL BLOCK SAMPLED ?
NO
YES
S402
SIZE < THRESHOLD ?
YES
"CHAR"
NO
S403
PORTRAIT SIZE OR LANDSCAPE SIZE ?
YES
NO
S404
SLANTED LINE ?
YES
"LINE"
NO
S405
SAMPLE WHITE PIXEL BLOCK SURROUNDED BY 4 CONNECTED LINES
S406
RECTANGULAR OUTLINE ?
YES
NO
S407
ARRANGEMENT OF WHITE PIXEL BLOCK IS GOOD ?
YES
"TABLE"
NO
"NONCHAR"
S408
SAMPLE BLACK PIXEL BLOCK FROM "NONCHAR" AND "TABLE"
END

FIG. 5

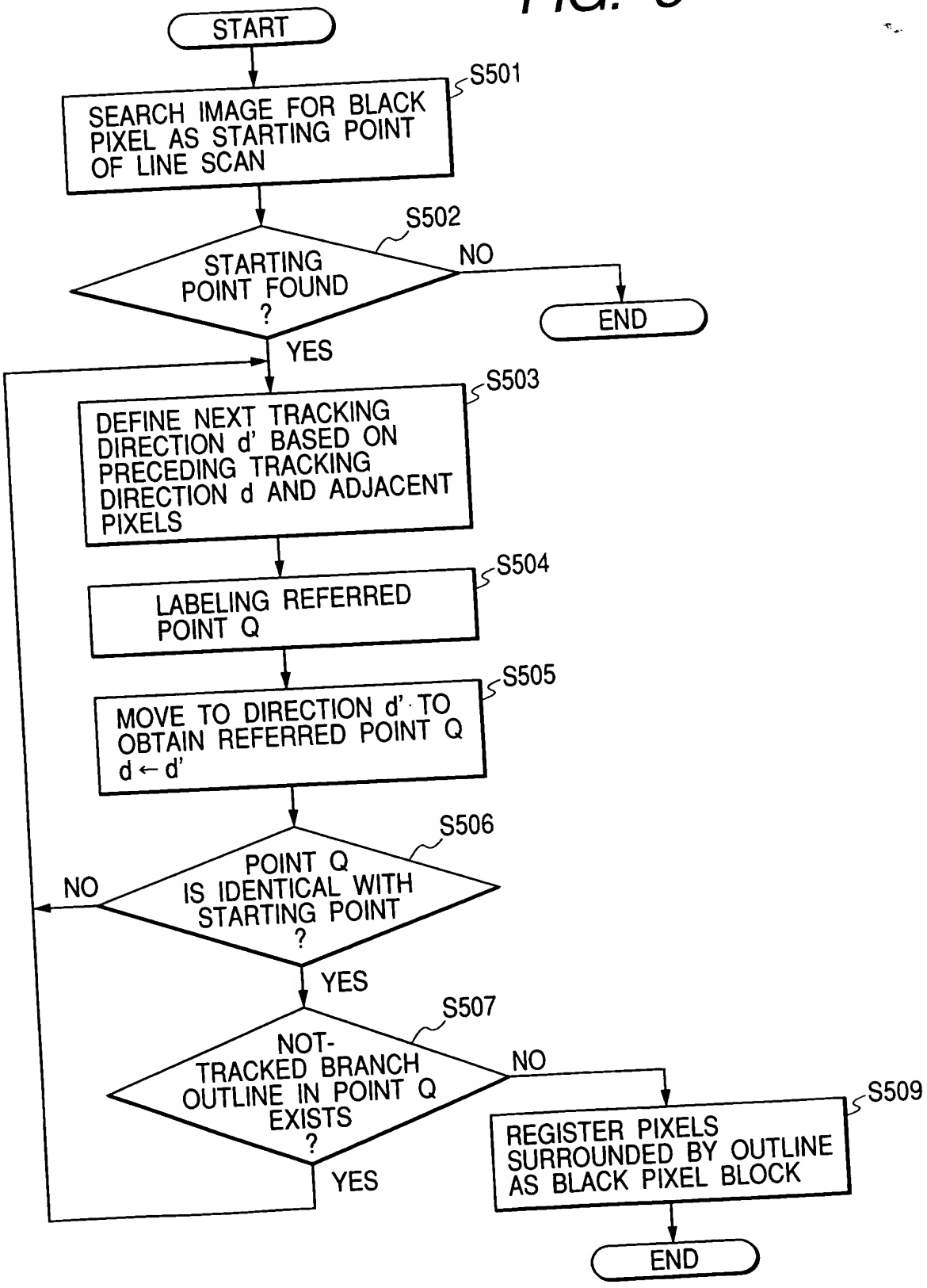

START
SEARCH IMAGE FOR BLACK PIXEL AS STARTING POINT OF LINE SCAN
S501
STARTING POINT FOUND ?
S502
NO
END
YES
DEFINE NEXT TRACKING DIRECTION d' BASED ON PRECEDING TRACKING DIRECTION d AND ADJACENT PIXELS
S503
LABELING REFERRED POINT Q
S504
MOVE TO DIRECTION d' TO OBTAIN REFERRED POINT Q
d ← d'
S505
POINT Q IS IDENTICAL WITH STARTING POINT ?
S506
NO
YES
NOT-TRACKED BRANCH OUTLINE IN POINT Q EXISTS ?
S507
NO
YES
REGISTER PIXELS SURROUNDED BY OUTLINE AS BLACK PIXEL BLOCK
S509
END

FIG. 6

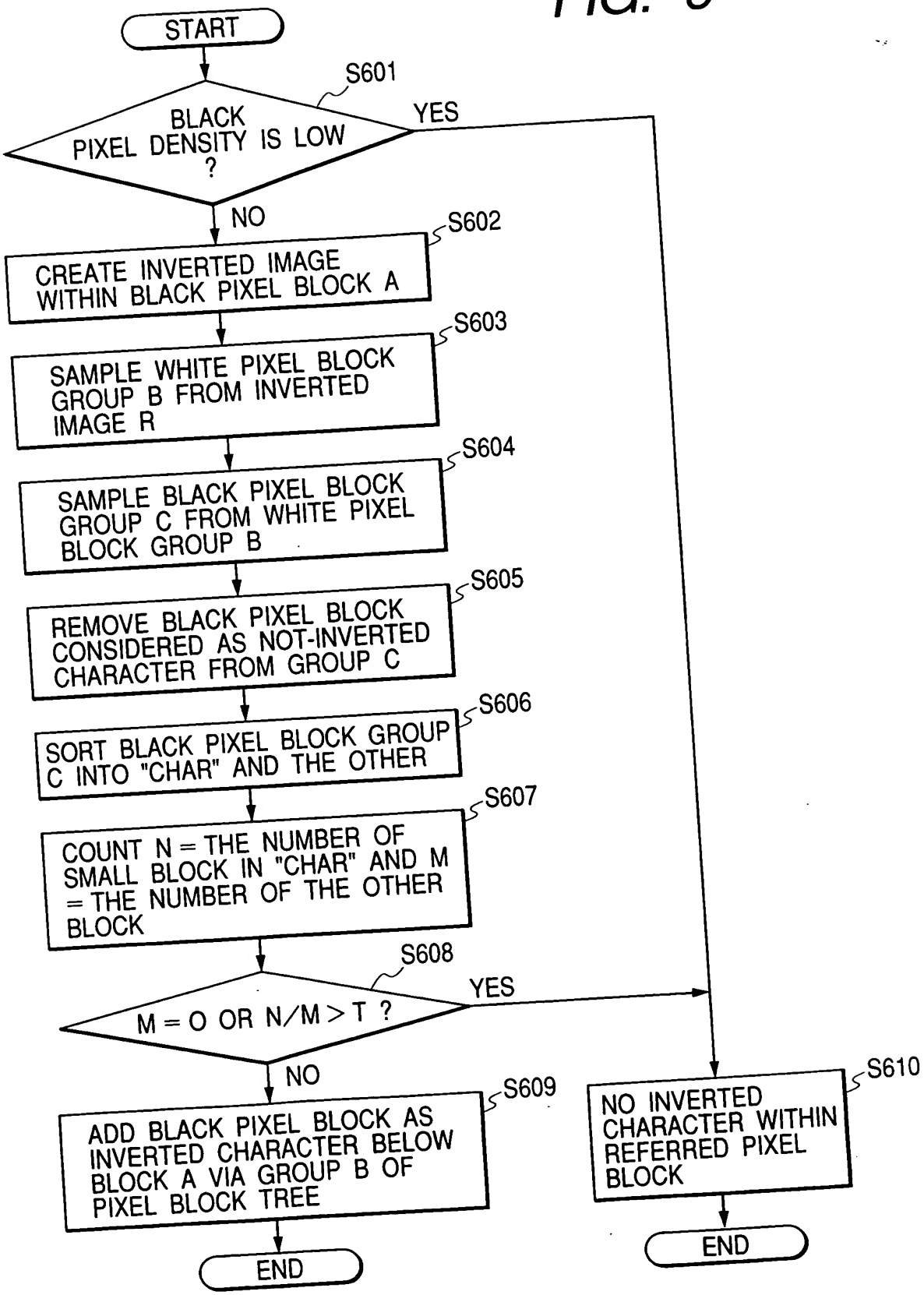

START
S601
BLACK PIXEL DENSITY IS LOW ?
YES
NO
S602
CREATE INVERTED IMAGE WITHIN BLACK PIXEL BLOCK A
S603
SAMPLE WHITE PIXEL BLOCK GROUP B FROM INVERTED IMAGE R
S604
SAMPLE BLACK PIXEL BLOCK GROUP C FROM WHITE PIXEL BLOCK GROUP B
S605
REMOVE BLACK PIXEL BLOCK CONSIDERED AS NOT-INVERTED CHARACTER FROM GROUP C
S606
SORT BLACK PIXEL BLOCK GROUP C INTO "CHAR" AND THE OTHER
S607
COUNT N = THE NUMBER OF SMALL BLOCK IN "CHAR" AND M = THE NUMBER OF THE OTHER BLOCK
S608
M = O OR N/M > T ?
YES
NO
S609
ADD BLACK PIXEL BLOCK AS INVERTED CHARACTER BELOW BLOCK A VIA GROUP B OF PIXEL BLOCK TREE
END
S610
NO INVERTED CHARACTER WITHIN REFERRED PIXEL BLOCK
END

FIG. 7

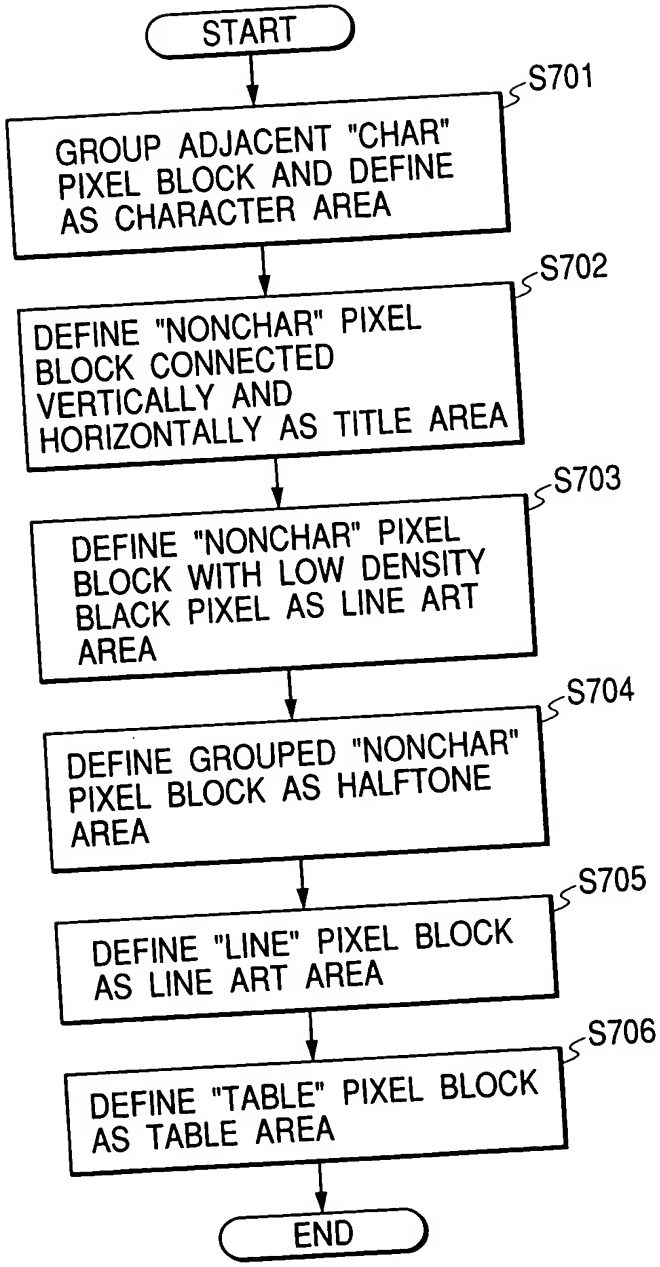

START
S701
GROUP ADJACENT "CHAR" PIXEL BLOCK AND DEFINE AS CHARACTER AREA
S702
DEFINE "NONCHAR" PIXEL BLOCK CONNECTED VERTICALLY AND HORIZONTALLY AS TITLE AREA
S703
DEFINE "NONCHAR" PIXEL BLOCK WITH LOW DENSITY BLACK PIXEL AS LINE ART AREA
S704
DEFINE GROUPED "NONCHAR" PIXEL BLOCK AS HALFTONE AREA
S705
DEFINE "LINE" PIXEL BLOCK AS LINE ART AREA
S706
DEFINE "TABLE" PIXEL BLOCK AS TABLE AREA
END

FIG. 8

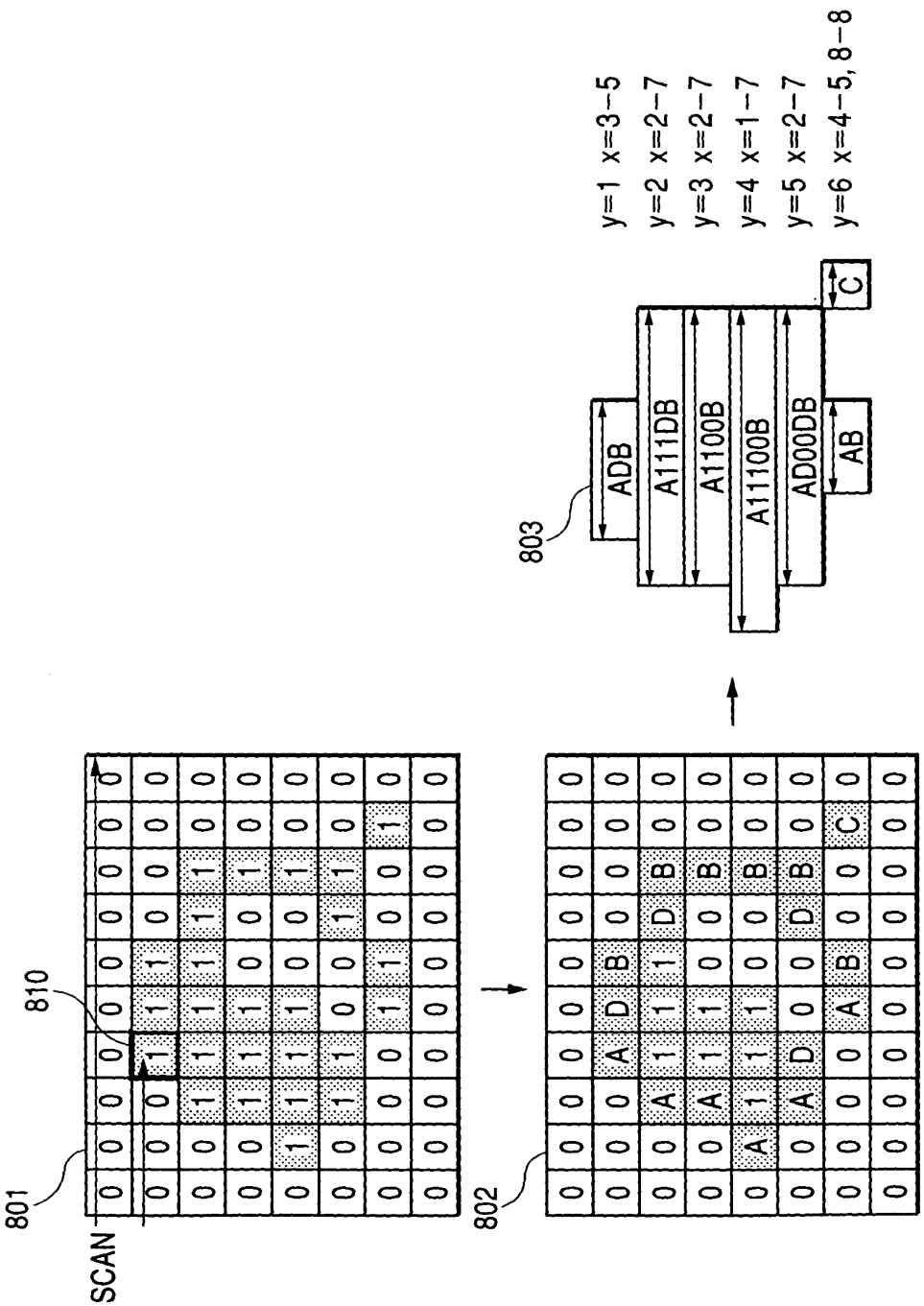

801
810
SCAN
802
803
ADB
A111DB
A1100B
A11100B
AD00DB
AB
C
y=1 x=3−5
y=2 x=2−7
y=3 x=2−7
y=4 x=1−7
y=5 x=2−7
y=6 x=4−5, 8−8

FIG. 9

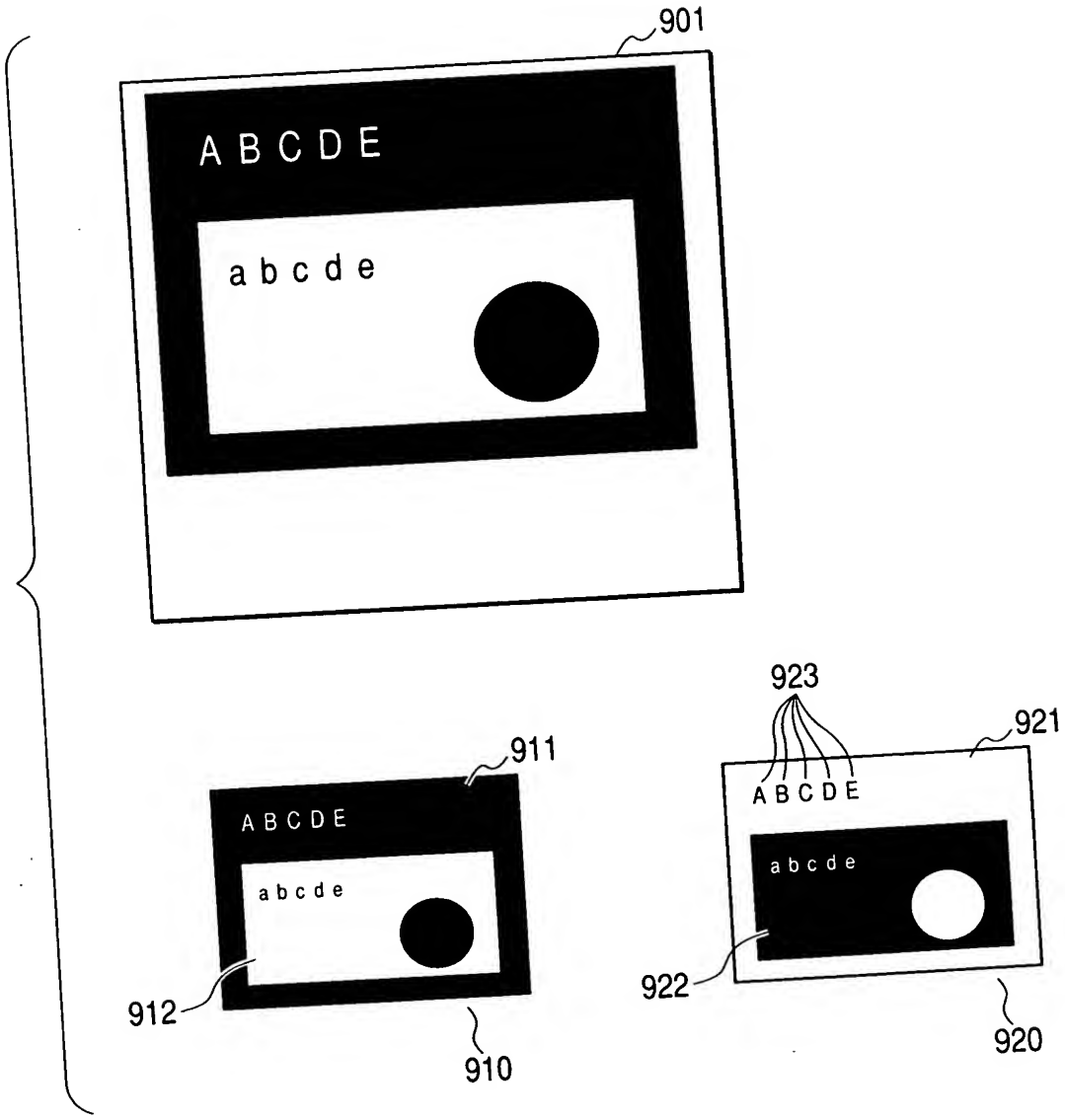

901
A B C D E
a b c d e
911
A B C D E
a b c d e
912
910
923
921
A B C D E
a b c d e
922
920

FIG. 10

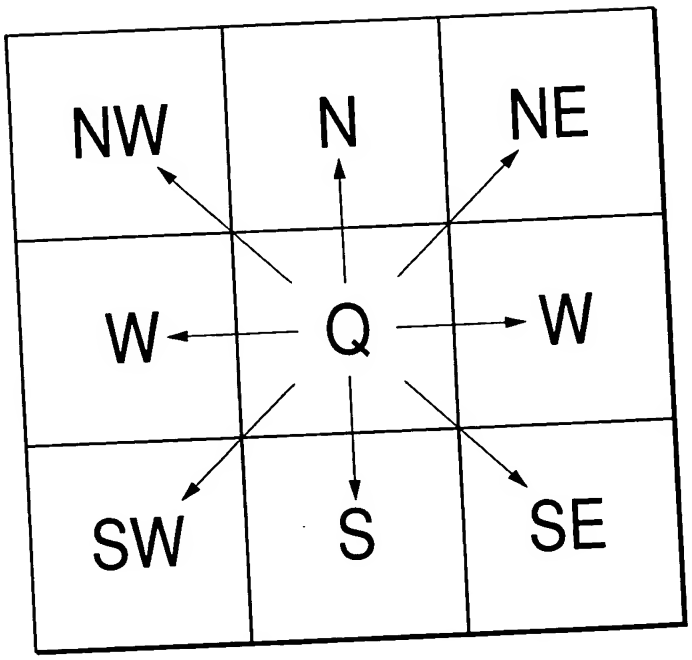


FIG. 11

d \ n	1	2	3	4	5	6	7
NW	NE	N	NW	W	SW	S	SE
W	NW	W	SW	S	SE	E	
SW	NW	W	SW	S	SE	E	NE
S	SW	S	SE	E	NE	N	
SE	SW	S	SE	E	NE	N	NW
E	SE	E	NE	N	NW	W	
NE	SE	E	NE	N	NW	W	SW
N	NE	N	NW	W	SW	S	

FIG. 12

d \ n	1	2	3	4	5	6	7
NW	*1	*1	*1	*1	*2	*2	*2
W	*4	*4	*3	*3	*3	*3	
SW	*4	*4	*3	*3	*3	*3	*2
S	*3	*3	*3	*3	*2	*2	
SE	*3	*3	*3	*3	*2	*2	*2
E	*4	*4	*1	*1	*1	*1	
NE	*4	*4	*1	*1	*1	*1	*2
N	*1	*1	*1	*1	*2	*2	

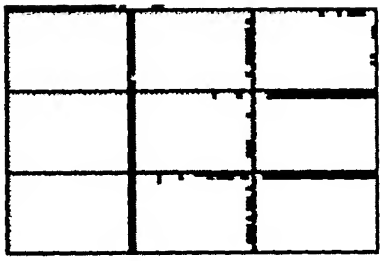
*1: LABEL C IS APPLIED WHEN LABEL A IS ALREADY APPLIED
IN OTHER CASE, LABEL B IS APPLIED

*2: LABEL C IS APPLIED

*3: LABEL C IS APPLIED WHEN LABEL C IS ALREADY APPLIED
IN OTHER CASE, LABEL A IS APPLIED

*4: LABEL D IS APPLIED

FIG. 13



EXAMPLE OF RECTANGULAR BLACK PIXEL BLOCK (LEFT)
AND NON-RECTANGULAR BLACK PIXEL BLOCK (RIGHT)

FIG. 14A

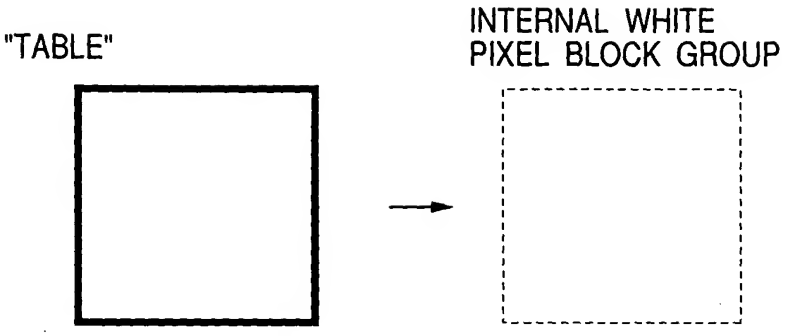


FIG. 14B

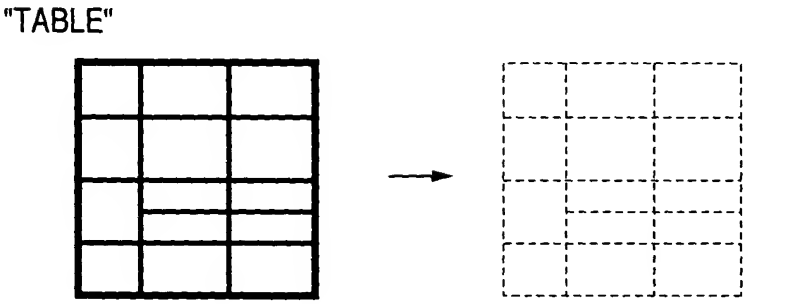


FIG. 14C

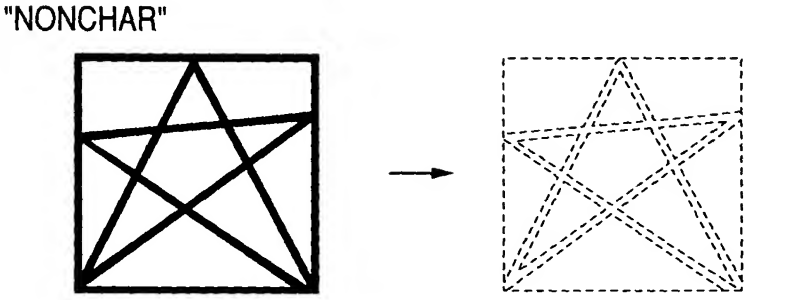


FIG. 15

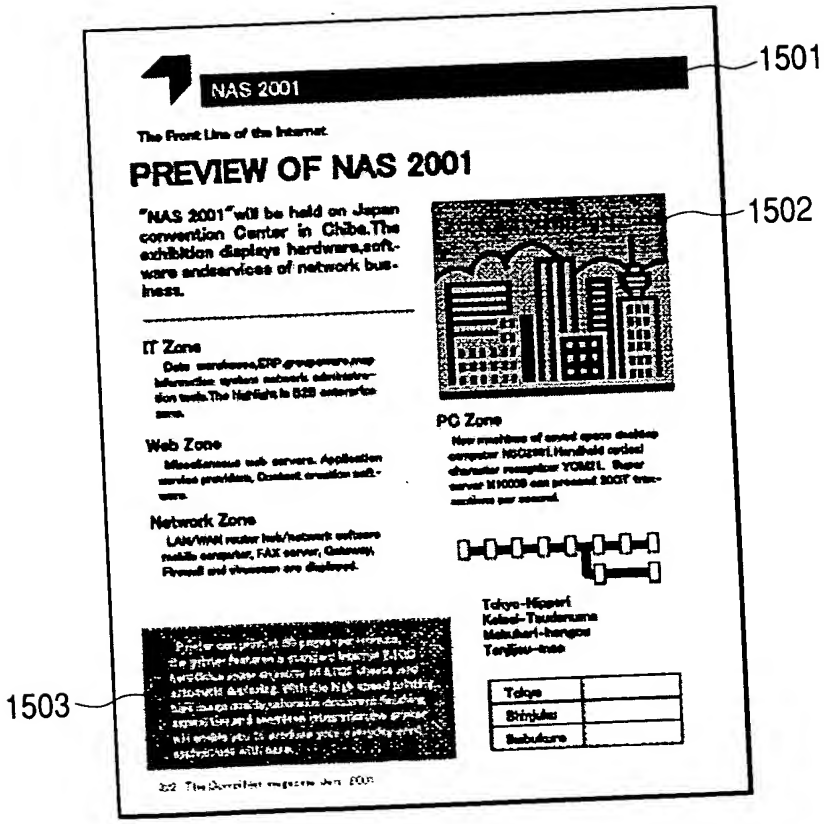

NAS 2001
1501
The Front Line of the Internet
PREVIEW OF NAS 2001
"NAS 2001"will be held on Japan convention Center in Chiba.The exhibition displays hardware,software andservices of network business.
1502
IT Zone
Web Zone
Network Zone
PC Zone
Tokyo-Higashi
Keisei-Tsudanuma
Tenjiku-machi
Tokyo
Shinjuku
Sabukuro
1503

FIG. 16

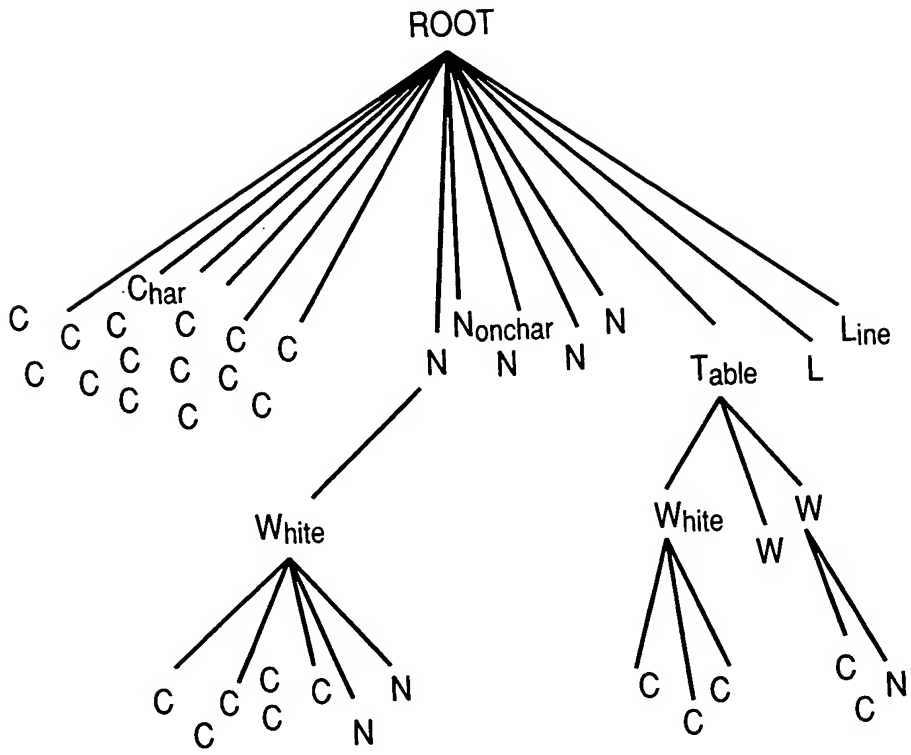


FIG. 17

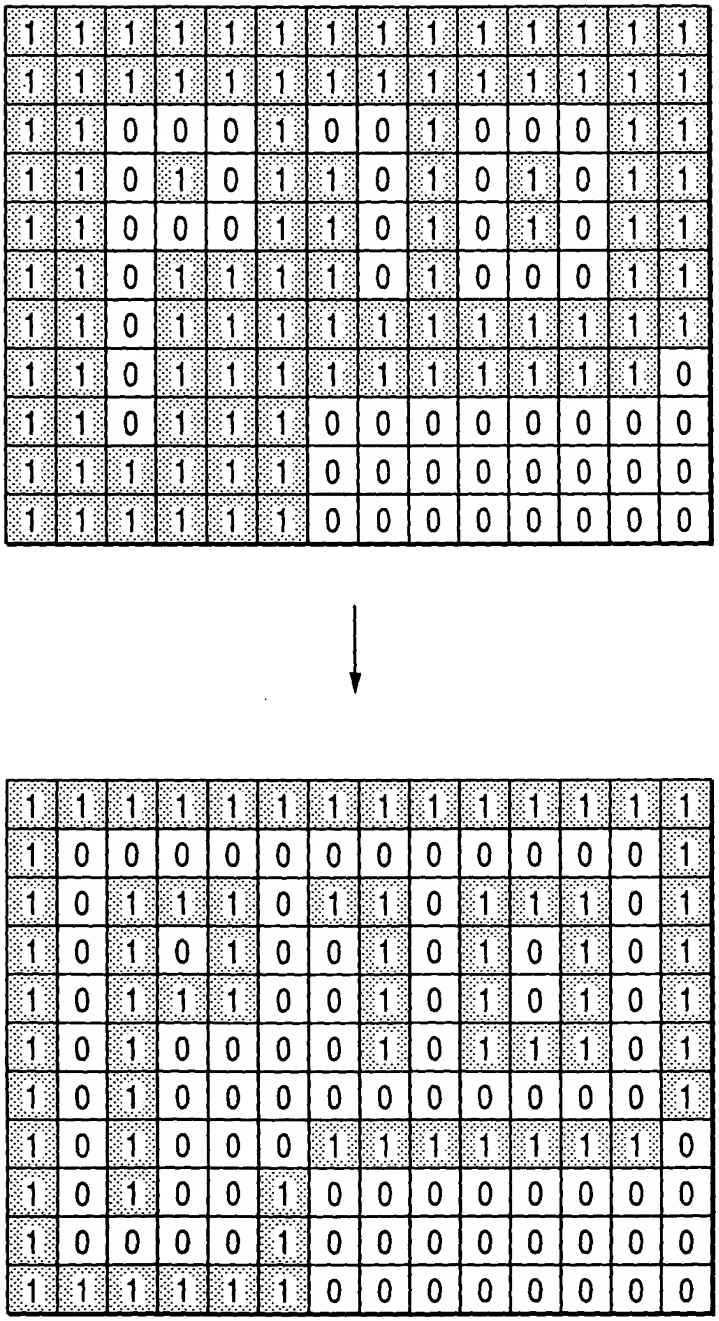

1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 0 0 0 1 0 0 1 0 0 0 1 1
1 1 0 1 0 1 1 0 1 0 1 0 1 1
1 1 0 0 0 1 1 0 1 0 1 0 1 1
1 1 0 1 1 1 1 0 1 0 0 0 1 1
1 1 0 1 1 1 1 1 1 1 1 1 1 1
1 1 0 1 1 1 1 1 1 1 1 1 1 0
1 1 0 1 1 1 0 0 0 0 0 0 0 0
1 1 1 1 1 1 0 0 0 0 0 0 0 0
1 1 1 1 1 1 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 0 0 0 0 0 0 0 0 0 0 0 0 1
1 0 1 1 1 0 1 1 0 1 1 1 0 1
1 0 1 0 1 0 0 1 0 1 0 1 0 1
1 0 1 1 1 0 0 1 0 1 0 1 0 1
1 0 1 0 0 0 0 1 0 1 1 1 0 1
1 0 1 0 0 0 0 0 0 0 0 0 0 1
1 0 1 0 0 0 1 1 1 1 1 1 1 0
1 0 1 0 0 1 0 0 0 0 0 0 0 0
1 0 0 0 0 1 0 0 0 0 0 0 0 0
1 1 1 1 1 1 0 0 0 0 0 0 0 0

FIG. 18

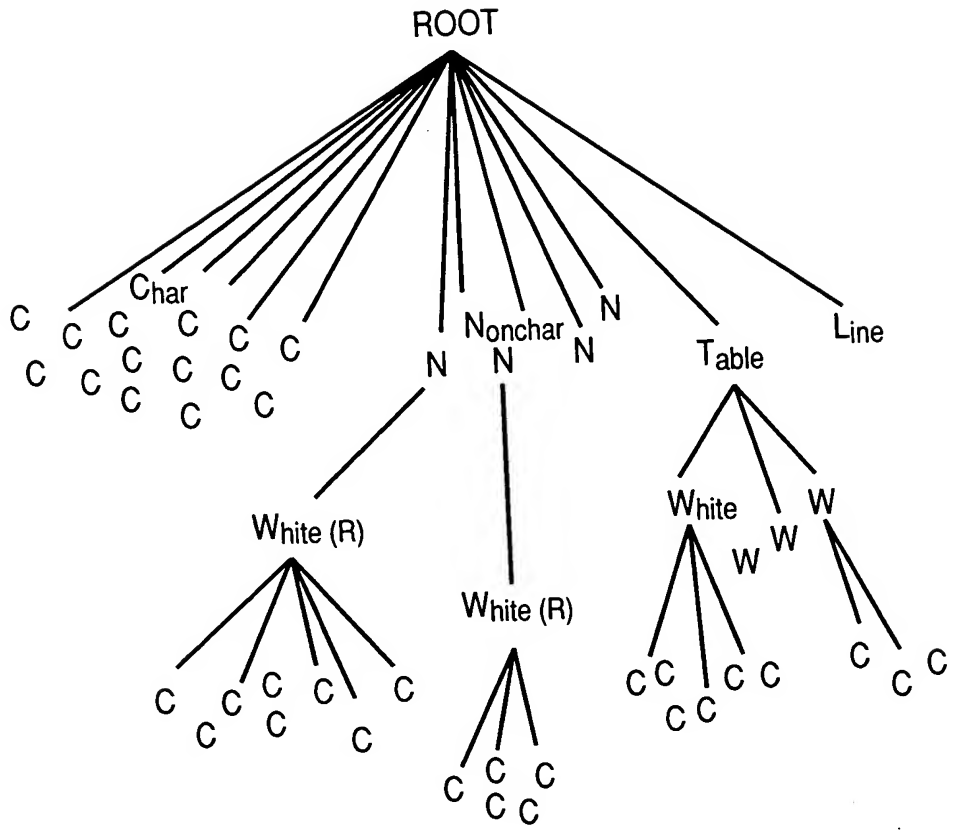


FIG. 19

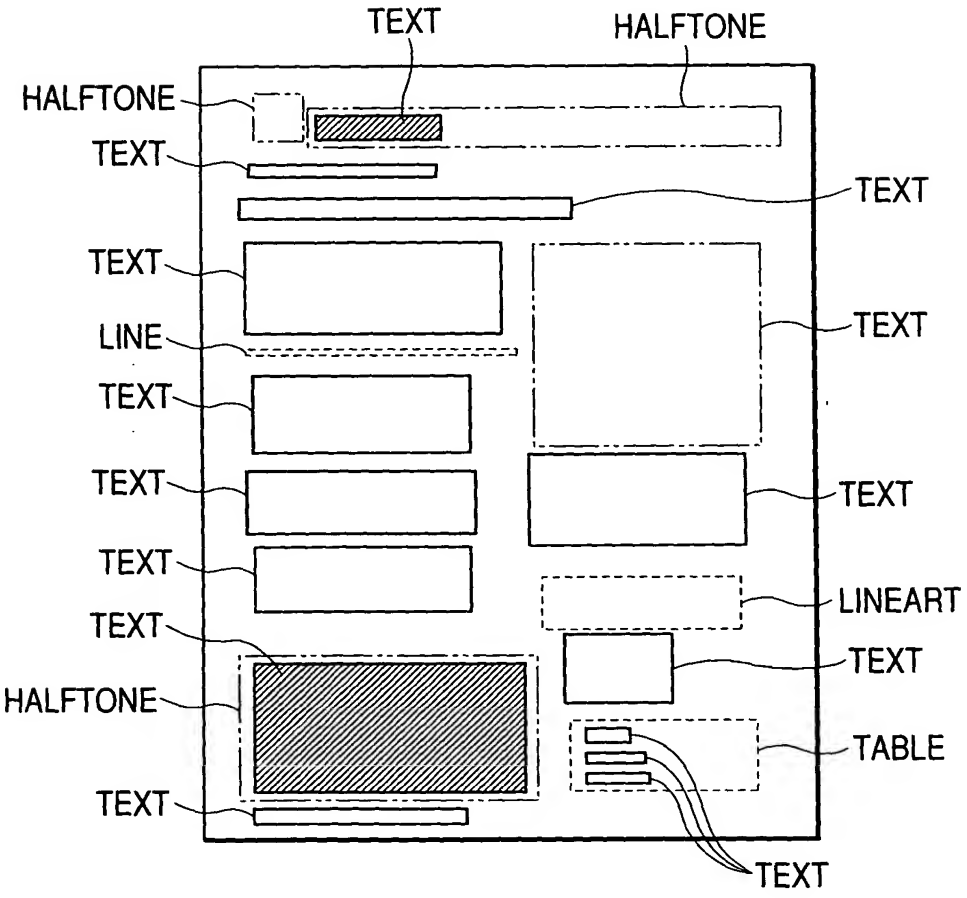

TEXT
HALFTONE
HALFTONE
TEXT
TEXT
TEXT
TEXT
LINE
TEXT
TEXT
TEXT
TEXT
TEXT
LINEART
TEXT
TEXT
HALFTONE
TABLE
TEXT
TEXT

FIG. 20

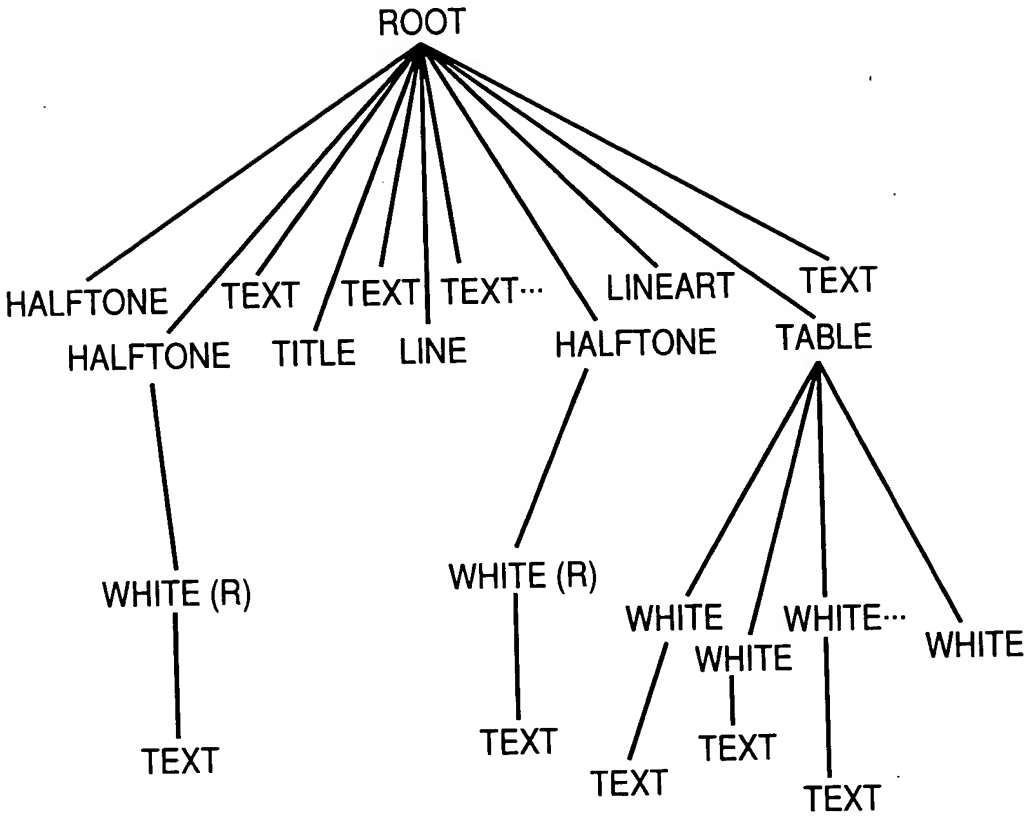

ROOT
HALFTONE
HALFTONE
TEXT
TITLE
TEXT
LINE
TEXT···
LINEART
HALFTONE
TEXT
TABLE
WHITE (R)
TEXT
WHITE (R)
TEXT
WHITE
WHITE
WHITE···
WHITE
TEXT
TEXT
TEXT

FIG. 21

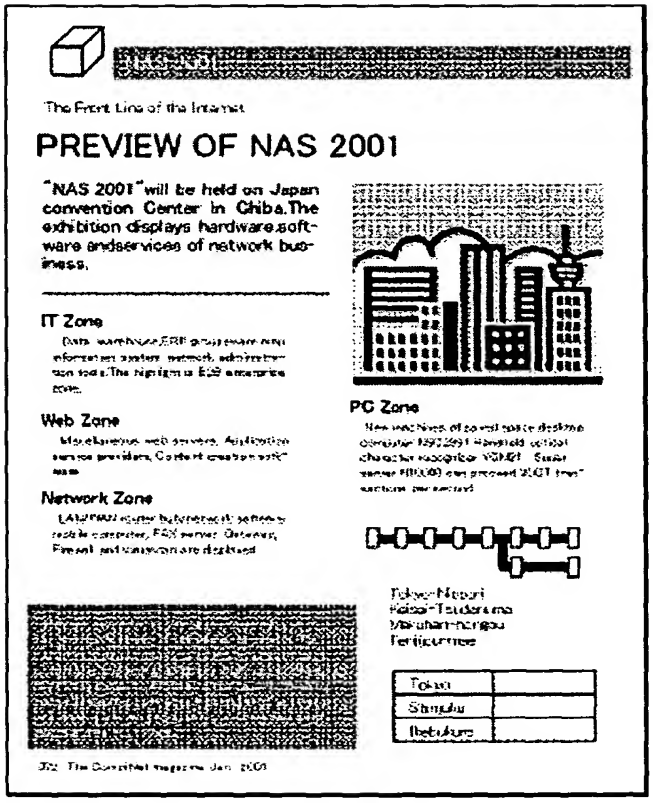

The Front Line of the Internet
PREVIEW OF NAS 2001
"NAS 2001" will be held on Japan convention Center in Chiba.The exhibition displays hardware,software and services of network business.
IT Zone
Web Zone
Network Zone
PC Zone
Tokyo
Shinjuku
Ikebukuro

FIG. 22

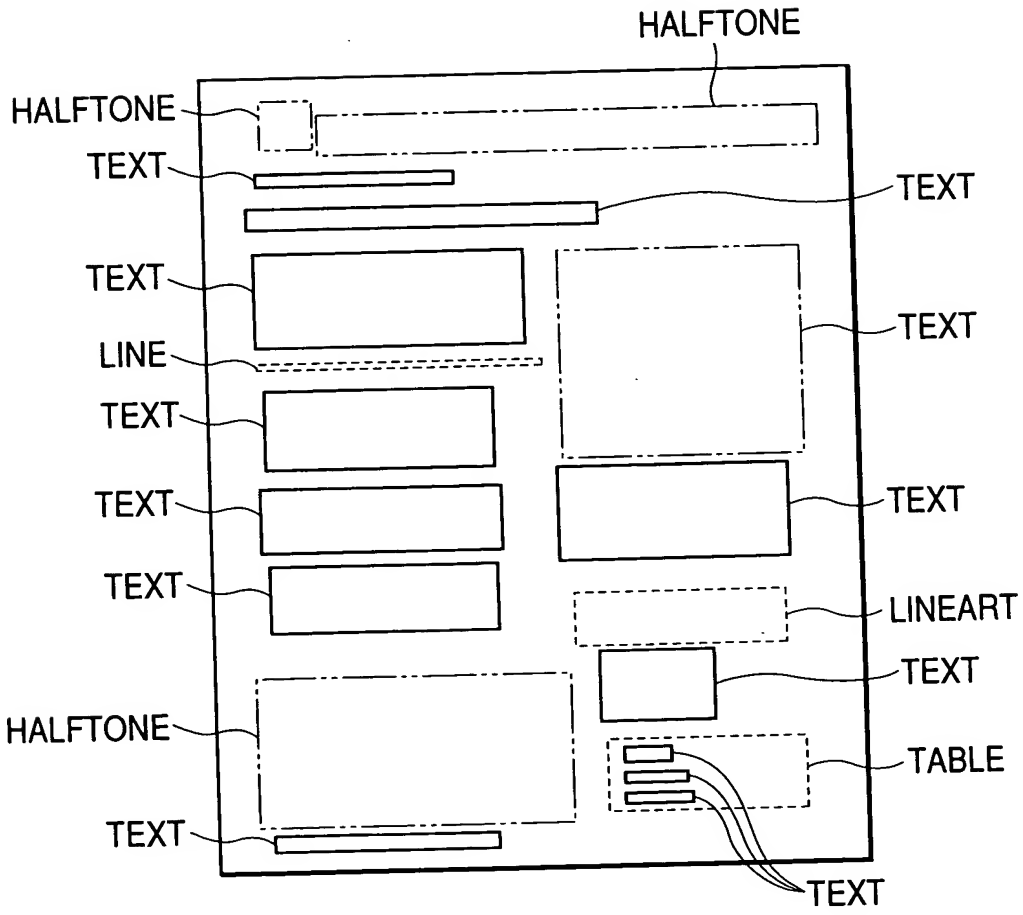
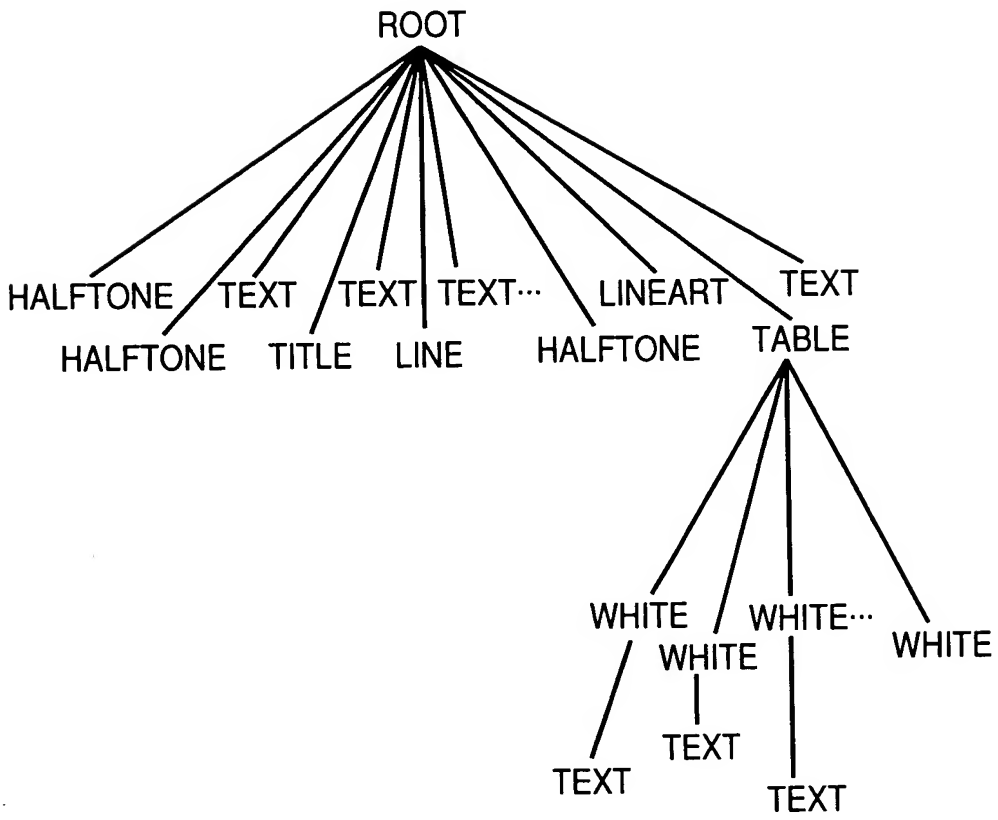

HALFTONE
HALFTONE
TEXT
TEXT
TEXT
TEXT
LINE
TEXT
TEXT
TEXT
TEXT
LINEART
TEXT
HALFTONE
TABLE
TEXT
TEXT

FIG. 23


ROOT
HALFTONE
HALFTONE
TEXT
TITLE
TEXT
LINE
TEXT···
LINEART
HALFTONE
TEXT
TABLE
WHITE
TEXT
WHITE
TEXT
WHITE···
TEXT
WHITE